On the Interaction between Gateway Scheduling and Routing*

Ibrahim Matta, A. Udaya Shankar

Department of Computer Science
University of Maryland
College Park, MD 20742

Abstract

Future computer networks are expected to provide different types of service. For this purpose, new algorithms and protocols have been proposed for gateway scheduling, flow control, and routing. The interaction between these three components is crucial to the performance of the network. Existing work has studied only the interaction between scheduling and flow control, assuming static routing. In this paper, we investigate the interaction between scheduling and adaptive routing. We view the network as a dynamical system. We apply the Liapunov direct method to derive stability conditions for the routes of different traffic classes. We show how with scheduling support for routing, the routes of the traffic classes can be isolated, thereby improving the overall network performance.

1 Introduction

In a packet-switched computer network, there are three components that are critical to providing different types of service: gateway scheduling discipline, flow control, and routing. The interaction between these three components is crucial to the performance of the network. Existing work has studied the interaction between scheduling and flow control, assuming static routing (e.g. [3, 4]). In this paper, we investigate the interaction between scheduling and adaptive routing.

We consider a network supporting two traffic classes: *delay-sensitive traffic*, and *throughput-sensitive traffic*. We consider two scheduling disciplines: the common FCFS (first-come-first-served), and one we refer to as TOSQ (type-of-service queueing). In TOSQ, each traffic class is assigned to a separate queue at the output link. The link bandwidth is allocated *equally* between the two queues in a round-robin fashion. Routes for each traffic class are computed using some link cost. The link cost for throughput-sensitive traffic is the measured utilization of the link. The link cost for delay-sensitive traffic depends on the link scheduling discipline. With FCFS, it is the measured link delay. With TOSQ, it is the measured delay that is experienced by delay-sensitive packets *only*.

This type-of-service network was recently studied in [7] via discrete-event simulation (a subset of the NSFNET-T1-Backbone topology was considered). It was shown that routing with TOSQ provides lower delays than routing with FCFS in a typical situation where the proportion of delay-sensitive traffic is small compared to the throughput-sensitive traffic [5].

The explanation given there is as follows: It is well known that delay-based routing does not perform well at high load when queueing delay is a significant part of measured link delay, which consists of queueing, transmission, and propagation delays [6]. Because delay-sensitive packets have a lower queueing delay under TOSQ, delay-based routing of delay-sensitive packets performs better and results in more stable routes for that traffic class. Meanwhile, the link utilization metric makes the throughput-sensitive traffic move away from the delay-sensitive traffic, taking the under-utilized routes. This has the effect of *isolating* the two traffic classes, resulting in a better overall network performance.

In this paper, we investigate this behavior analytically. We consider a simplified situation consisting of a source node sending traffic to a destination node along two paths. The first path represents low delay routes, and the second path represents high capacity routes.

*This work is supported in part by ONR and DARPA under contract N00014-91-C-0195 to Honeywell and Computer Science Department at the University of Maryland, and by National Science Foundation Grant No. NCR 89-04590. The views, opinions, and/or findings contained in this paper are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, ONR, the U.S. Government or Honeywell. Computer facilities were provided in part by NSF grant CCR-8811954.

We model this network as a dynamical system [1]. We represent isolation by a *stable* state where *every* delay-sensitive connection stays on the first path, and *every* throughput-sensitive connection stays on the second path.

Intuitively, isolation is desirable since otherwise it becomes more likely that both traffic classes will move away from a highly loaded link (i.e. with high delay and utilization) at the same routing update. Such simultaneous traffic shifts degrade the overall performance, and in particular result in a higher delay and network under-utilization.

Using a worst-case analysis, we first obtain sufficient conditions on the system parameters to reach isolation, independent of the starting state. We then apply the Liapunov direct method [9] to obtain more refined sufficient conditions for isolation as a function of the starting state; i.e. these conditions define a *domain of attraction* to isolation. (In a real network, the starting state would be the result of arrivals of new connections, departures of old connections, failure/recovery of links, etc.) Under certain parameters' values, the domain of attraction for routing with TOSQ is shown to be larger than that for routing with FCFS. Our monte-carlo simulations also show that routing with TOSQ converges to isolation much faster than routing with FCFS.

The rest of the paper is organized as follows. In Section 2, we give the model. In Section 3, we give the stability conditions obtained using worst-case analysis. Section 4 contains our stability analysis using the Liapunov direct method. Monte-carlo simulation results are presented in Section 5. We conclude the paper in Section 6. Details of derivations and proofs are not presented here for space reasons, and can be found in [8].

2 Model

We model a network by a node S sending traffic to node D along two paths. Path 1 represents low delay routes, and path 2 represents high capacity routes. Path i ($i = 1, 2$) has propagation delay P_i time units and average transmission capacity C_i packets/time unit. There are N delay-sensitive connections, and M throughput-sensitive connections from S to D. For every connection, packets originate at S according to a Poisson process, and without loss of generality we assume an arrival rate of 1 packet/time unit.

At any instant, we describe the state of the network by the tuple (x, y), where x is the number of delay-sensitive connections on path 1, and y is the number of throughput-sensitive connections on path 1. To model routing updates, we use a discrete-time flow approach as in [2]. We assume that (some or all) connections periodically update their routes to D every $\triangle$ time units, where $\triangle$ is long enough for the network to reach steady-state after a routing update. Routes, and hence the network state, are updated at discrete time instants $(k+1)\triangle$, $k = 0, 1, 2, \cdots$. Let (x_k, y_k) be the network state immediately after time $k\triangle$. At an update instant $(k+1)\triangle$, we use steady-state $M/M/1$ results to estimate link costs based on (x_k, y_k). Using these link costs, routes are updated and consequently the new network state (x_{k+1}, y_{k+1}) is obtained. We denote by $T_{i,k+1}$ and $\rho_{i,k+1}$ the delay and utilization cost of path i, respectively, at time $(k+1)\triangle$.

With a FCFS scheduling discipline at S, we can write the delay link costs as follows:

$$\begin{aligned} T_{1,k+1} &= \frac{1}{C_1 - (x_k + y_k)} + P_1 \\ T_{2,k+1} &= \frac{1}{C_2 - (\bar{x}_k + \bar{y}_k)} + P_2 \end{aligned} \tag{1}$$

where $\bar{x}_k = N - x_k$ and $\bar{y}_k = M - y_k$ denote the number of delay-sensitive connections and the number of throughput-sensitive connections, respectively, on path 2.

The utilization link costs are

$$\begin{aligned} \rho_{1,k+1} &= \frac{x_k + y_k}{C_1} \\ \rho_{2,k+1} &= \frac{\bar{x}_k + \bar{y}_k}{C_2} \end{aligned} \tag{2}$$

The network state is updated using the costs of the two paths as follows:

$$\begin{aligned} x_{k+1} &= \begin{cases} x_k + \alpha_k\, \bar{x}_k & \text{if } T_{1,k+1} \leq T_{2,k+1} \\ (1-\alpha_k)\, x_k & \text{otherwise} \end{cases} \\ y_{k+1} &= \begin{cases} y_k + \alpha_k\, \bar{y}_k & \text{if } \rho_{1,k+1} < \rho_{2,k+1} \\ (1-\alpha_k)\, y_k & \text{otherwise} \end{cases} \end{aligned} \tag{3}$$

The parameter α_k ($0 < \alpha_k \leq 1$) reflects the amount of traffic rerouted. It can also be thought of as the degree of routing update synchronization at different nodes. We assume that α_k is uniformly distributed over $[\alpha_{MIN}, \alpha_{MAX}]$, where $\alpha_{MAX} - \alpha_{MIN} = 0.2$, and $0.1 < \frac{\alpha_{MAX}+\alpha_{MIN}}{2} \leq 0.9$.

With a TOSQ scheduling discipline at S, the two queues at an output link are correlated, which makes the analysis difficult. A number of approximate solutions for such systems have been proposed. One common approach is to approximate the system by two loosely-coupled $M/M/1$ queues [10]. The service rate of each queue depends on the utilization of the other queue.

Define $C_{i,k}^{eff}$ as the effective capacity available for the delay-sensitive traffic on path i after time $k\triangle$. We have

$$C_{i,k}^{eff} \geq 0.5C_i \tag{4}$$

with the worst case occurring when the other queue is *always* not empty. Assuming each queue is $M/M/1$, we obtain the following:

$$\begin{array}{rcl} C_{1,k}^{eff} & = & \frac{(C_1-0.5(y_k-x_k))+\sqrt{(C_1-0.5(y_k-x_k))^2-2C_1x_k}}{2} \\ C_{2,k}^{eff} & = & \frac{(C_2-0.5(\bar{y}_k-\bar{x}_k))+\sqrt{(C_2-0.5(\bar{y}_k-\bar{x}_k))^2-2C_2\bar{x}_k}}{2} \end{array} \tag{5}$$

Thus we have the following delay link costs with TOSQ scheduling:

$$\begin{array}{rcl} T_{1,k+1} & = & \dfrac{1}{C_{1,k}^{eff}-x_k}+P_1 \\ T_{2,k+1} & = & \dfrac{1}{C_{2,k}^{eff}-\bar{x}_k}+P_2 \end{array} \tag{6}$$

The utilization link costs are as defined in (2). The network state is updated as in (3).

We refer to the iteration defined in (3) as I, i.e. $(x_{k+1}, y_{k+1}) = I(x_k, y_k)$. I is a mapping from a set G into itself, where $G = \{(x, y) : 0 \leq x \leq N \wedge 0 \leq y \leq M\}$. The sequence of points (x_1, y_1), (x_2, y_2), $\ldots$ is called the *trajectory* of the system [9]. The trajectory may or may not converge to a fixed point (x^*, y^*), i.e. $(x^*, y^*) = I(x^*, y^*)$. *Isolation* is equivalent to say that I converges to the fixed point $(N, 0)$. In the following two sections, we give sufficient conditions for isolation.

3 Worst-case Analysis

In this section, we present sufficient conditions for isolation regardless of the starting state and the values of α_k. Figure 1 depicts the *load regions*, i.e. values of (M, N), defined by the sufficient conditions.

Observe that as long as N is less than $0.5C_1 - \frac{1}{(\frac{1}{C_2}+P_2-P_1)}$, routing with TOSQ provides isolation (regardless of the starting state and the values of α_k) for higher values of M than with FCFS. This effect becomes more pronounced as C_2 increases.

4 Liapunov Stability Analysis

We use the Liapunov direct method [9] to obtain sufficient conditions for isolation as a function of the starting state and the values of α_k. Thus, these conditions capture more information; i.e. there can be a system configuration that does not satisfy the sufficient conditions of Section 3, but still reaches isolation for some starting states or values of α_k.

For ease of presentation, we assume zero propagation delays, i.e. $P_1 = P_2 = 0$. Also, it is more convenient to deal with the fixed point $(0, 0)$ rather than

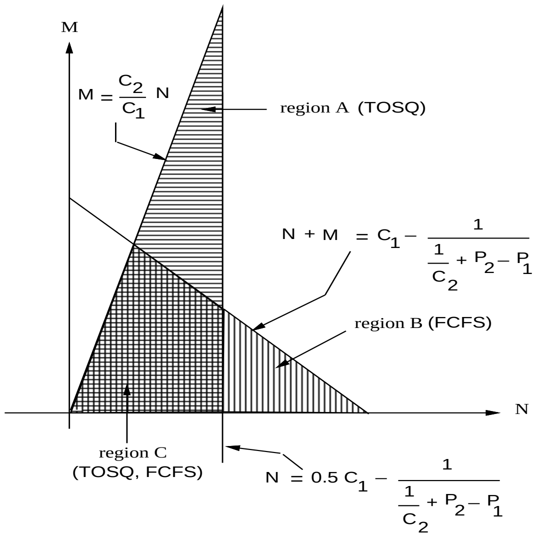

Figure 1: Load regions satisfying sufficient conditions for isolation.

$(N, 0)$. Thus, we define the network state by $(\bar{x}, y)$ instead of (x, y). Then, the system behavior with FCFS in place is described by the following:

$$\begin{array}{rcl} \bar{x}_{k+1} & = & (1-\alpha_k)\,\bar{x}_k + \alpha_k\, N\, \delta_k \\ y_{k+1} & = & (1-\alpha_k)\, y_k + \alpha_k\, M\, \beta_k \end{array} \tag{7}$$

where

$$\begin{array}{rcl} \delta_k & = & \left\{ \begin{array}{ll} 1 & \text{if } \bar{x}_k - y_k < L_1 = \frac{C_2-C_1+N-M}{2} \\ 0 & \text{otherwise} \end{array} \right. \\ \beta_k & = & \left\{ \begin{array}{ll} 1 & \text{if } \bar{x}_k - y_k > L_2 = \frac{C_2N-C_1M}{C_1+C_2} \\ 0 & \text{otherwise} \end{array} \right. \end{array} \tag{8}$$

At the fixed point, $\bar{x}_k \to 0, \bar{x}_{k+1} \to 0, y_k \to 0$, and $y_{k+1} \to 0$. Consequently, for the equations (7) to be satisfied, $\delta_k \to 0$ and $\beta_k \to 0$, which imply that $L_1 \leq 0$ and $L_2 \geq 0$ are necessary (but not sufficient) conditions for convergence to the origin.

Define $D_1 = \{(\bar{x}, y) : L_1 \leq \bar{x} - y \leq L_2 \ \wedge\ 0 \leq \bar{x} \leq N \ \wedge\ 0 \leq y \leq M\}$. (See Figure 2.)

Define the scalar function $V(\bar{x}, y)$:

$$V(\bar{x}, y) = \bar{x}^2 + y^2 \tag{9}$$

We show that $V(\bar{x}, y)$ is a Liapunov function in D_1, which implies that starting from any point in D_1, the trajectory stays inside D_1, i.e., $\delta_k = \beta_k = 0$ for all $(\bar{x}_k, y_k)$ along the trajectory, approaching the origin as shown in Figure 2.

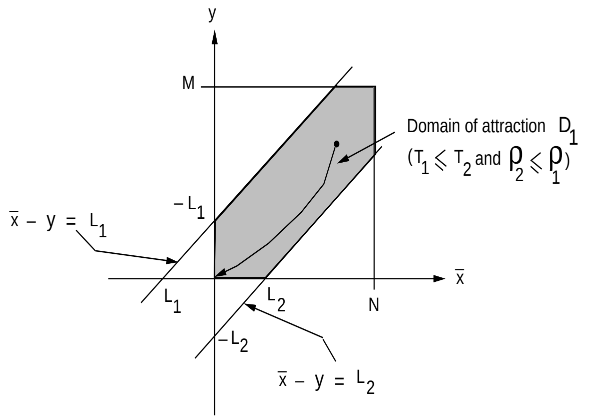

Figure 2: Domain of attraction for routing with FCFS with zero propagation delays.

Lemma 4.1 *$V(\bar{x}, y)$ is a Liapunov function in D_1. That is: (1) $V(\bar{x}, y)$ is positive definite. (2) For all $(\bar{x}_k, y_k) \in D_1 - \{(0,0)\}$, $V(\bar{x}_{k+1}, y_{k+1}) < V(\bar{x}_k, y_k)$ and $(\bar{x}_{k+1}, y_{k+1}) \in D_1$.*

From lemma 4.1 and the Liapunov stability theory [9], we have the following theorem:

Theorem 4.1 *For routing with FCFS scheduling, any starting state in D_1 (the shaded area in Figure 2) leads to the origin, regardless of the values of α_k.*

With TOSQ in place, the system behavior is described by the same difference equations (7) except that δ_k is defined as

$$\delta_k = \begin{cases} 1 & \text{if } \bar{x}_k < L_{3,k} = \frac{C_{2,k}^{eff} - C_{1,k}^{eff} + N}{2} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

At the fixed point, $x_k \to N, \bar{x}_k \to 0$, $y_k \to 0, \bar{y}_k \to M$. Then, from equations (5), we have $C_{2,k}^{eff} \to C_2 - 0.5M$, and $C_{1,k}^{eff} \to C_1$. Thus, $L_{3,k} \to L_3^*$, where $L_3^* = \frac{(C_2 - 0.5M) - C_1 + N}{2}$. For equations (7) to be satisfied at the fixed point, $\delta_k \to 0$ and $\beta_k \to 0$. Then $L_3^* \leq 0$ and $L_2 \geq 0$ are necessary conditions for convergence to the origin.

As we have done with FCFS, we want to show that $V(\bar{x}, y)$, defined in (9), is a Liapunov function in some region around the origin. Call this region D_2. The goal is to show that starting at any point in D_2, $\delta_k = \beta_k = 0$ for all $(\bar{x}_k, y_k)$ along the trajectory.

$\bar{x}_k \geq L_{3,k}$ implies $\delta_k = 0$. $\bar{x}_k - y_k \leq L_2$ implies $\beta_k = 0$. Because the expression $L_{3,k}$ is hard to work with, we try to find a simpler expression, say $L_{4,k}$, such that $L_{4,k} \geq L_{3,k}$ (then $\bar{x}_k \geq L_{4,k}$ implies $\delta_k = 0$). We would then define $D_2 = \{(\bar{x}, y) : \bar{x} \geq L_4 \ \wedge \ \bar{x} - y \leq L_2 \wedge 0 \leq \bar{x} \leq N \wedge 0 \leq y \leq M\}$, and attempt to show that $V(\bar{x}, y)$, defined in (9), is a Liapunov function in D_2.

To find $L_{4,k}$ greater than or equal $L_{3,k}$, we need an upper bound on $C_{2,k}^{eff}$, and a lower bound on $C_{1,k}^{eff}$. From equations (5), we see that $C_{2,k}^{eff} \leq C_2 - 0.5(\bar{y}_k - \bar{x}_k)$. We could not find an appropriate lower bound on $C_{1,k}^{eff}$. So we made the approximation $C_{1,k}^{eff} \approx C_1 - 0.5y_k$ (the accuracy of this is discussed below). This approximation leads to the following expression:

$$L_{4,k} = \frac{(C_2 - 0.5(\bar{y}_k - \bar{x}_k)) - (C_1 - 0.5y_k) + N}{2} \tag{11}$$

Noting that $\bar{x}_k \geq L_{4,k}$ implies $3\bar{x}_k - 2y_k \geq 4L_3^*$, we thus have $D_2 = \{(\bar{x}, y) : 3\bar{x} - 2y \geq 4L_3^* \ \wedge \ \bar{x} - y \leq L_2 \ \wedge \ 0 \leq \bar{x} \leq N \ \wedge \ 0 \leq y \leq M\}$. Figure 3 depicts this region.

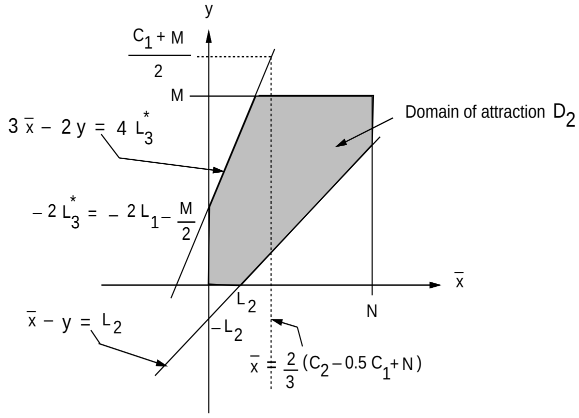

Figure 3: Domain of attraction for routing with TOSQ with zero propagation delays.

Theorem 4.2 *For routing with TOSQ scheduling, assuming $C_{1,k}^{eff} \approx C_1 - 0.5y_k$, any starting state in D_2 (the shaded area in Figure 3) leads to the origin, regardless of the values of α_k.*

The domains of attraction we have just found are not the largest, i.e. there may be points outside the domains which lead to the origin. This depends on the values of α_k. In particular, the domains are indeed largest for high enough values of α_k. On the other hand, they are not for small values of α_k.

Accounting for the propagation delays, and considering the fixed point $(N, 0)$ rather than $(0, 0)$, Figures 4 and 5 depict the domains of attraction for routing with FCFS and TOSQ, respectively, regardless of the values of α_k.

5 Monte-carlo Simulations

We use monte-carlo simulations to verify our assumption $C_{1,k}^{eff} \approx C_1 - 0.5 y_k$, which was made for routing with TOSQ scheduling. We also look at transient characteristics such as convergence time.

Figure 6 shows the domain of attraction corresponding to $(N, 0)$ for routing with TOSQ, for $C_1 = 30, C_2 = 50, P_1 = 0.5, P_2 = 1, N = 20, M = 30$, and $\alpha_k \ \epsilon \ [0.8, 1]$. Simulation results show fewer starting states leading to isolation than expected in the region where $x > 13$ ($x = 13$ is the value of x at which $\frac{1}{0.5C_1 - x} + P_1 = \frac{1}{(C_2 + 0.5\bar{x}) - \bar{x}} + P_2$). This is because the assumption $C_{1,k}^{eff} \approx C_1 - 0.5 y_k$ we made in our analysis is slightly optimistic; in fact, it can be shown that $C_{1,k}^{eff} \leq C_1 - 0.5 y_k$. Regardless of that, consistent with our analytic results, we observe that, for high values of α_k ($\alpha_{MIN} \geq 0.6$), isolation occurs for more starting states with TOSQ than with FCFS, in particular for higher values of y_0.

For low to moderate values of α_k ($\alpha_{MAX} \leq 0.6$), the system behavior becomes independent of the starting state (x_0, y_0). The system configurations for which isolation occurs are found to be the same with both FCFS and TOSQ [8]. We also observe that routing with TOSQ gives better transient performance, i.e. less oscillations and faster convergence to isolation.

6 Concluding Remarks

In this paper, we showed why in an integrated services environment, routing with some form of non-FCFS scheduling support can provide significant performance improvement. We are currently extending our analysis to obtain transient measures. Analyzing the interactions between adaptive routing and other scheduling algorithms is an open research area. Future work is also needed to explore the interaction between all components of the congestion control problem, namely scheduling, flow control, and routing, on arbitrary network topologies.

References

[1] K-H. Becker and M. Dorfler. *Dynamical Systems and Fractals.* Cambridge University Press, 1989.

[2] D. Bertsekas and R. Gallager. *Data Networks*, pages 297–333. Prentice-Hall, Inc., 1987.

[3] J. Bolot and A.U. Shankar. A Discrete-Time Stochastic Approach to Flow Control Dynamics. In Proc. *GLOBECOM '92*, Orlando, Florida, December 1992.

[4] A. Demers, S. Keshav, and S. Shenker. Analysis and Simulation of a Fair Queueing Algorithm. In Proc. *ACM SIGCOMM '89*, pages 1–12, Austin, Texas, September 1989.

[5] D. Johnson. NSFnet Report. In Proc. *of the Nineteenth Internet Engineering Task Force*, pages 377–382, University of Colorado, National Center for Atmospheric Research, December 1990.

[6] A. Khanna and J. Zinky. A Revised ARPANET Routing Metric. In Proc. *ACM SIGCOMM '89*, pages 45–56, September 1989.

[7] I. Matta and A.U. Shankar. Type-of-Service in Adaptive Next-Hop Routing. CS-TR-2963, University of Maryland, College Park, MD 20742, September 1992. Available by anonymous ftp at ftp.cs.umd.edu:pub/MaRS/Papers.

[8] I. Matta and A.U. Shankar. On the Interaction between Gateway Scheduling and Routing. CS-TR-3102, University of Maryland, College Park, MD 20742, July 1993. Available by anonymous ftp at ftp.cs.umd.edu:pub/MaRS/Papers.

[9] K. Ogata. *Discrete-Time Control Systems.* Prentice-Hall, Inc., 1987.

[10] W. Zhu and S. Chanson. Adaptive Threshold-based Scheduling for Real-Time and Non-Real-Time Traffic. In Proc. *IEEE RTSS '92*, pages 125–135, Phoenix, Arizona, December 1992.

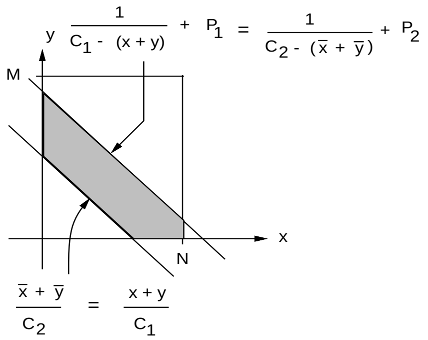

Figure 4: Domain of attraction for routing with FCFS, accounting for propagation delays.

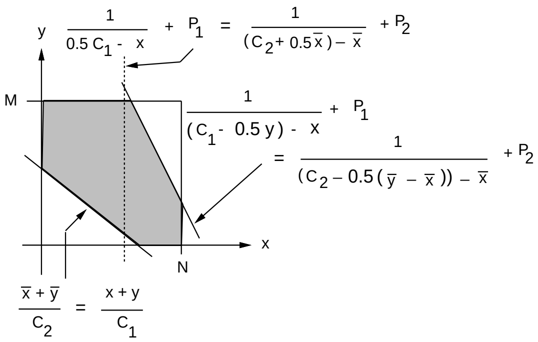

Figure 5: Domain of attraction for routing with TOSQ, accounting for propagation delays.

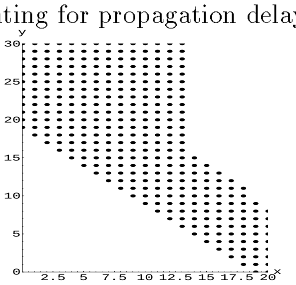

Figure 6: Domain of attraction for routing with TOSQ ($\alpha_k \ \epsilon \ [0.8, 1]$).